#2

RAW SEQUENCE LISTING   DATE: 01/24/2002
PATENT APPLICATION: US/09/892,287   TIME: 19:06:31

Input Set : **N:\Crf3\RULE60\09892287.raw**
Output Set: **N:\CRF3\01242002\I892287.raw**

SEQUENCE LISTING

```
  3  (1) GENERAL INFORMATION:
  5      (i) APPLICANT: Hillman, Jennifer L.
  6                     Lal, Preeti
  7                     Corley, Neil C.
  8                     Shah, Purvi
 10     (ii) TITLE OF INVENTION: NEW HUMAN PHOSPHATIDYLINOSITOL
 11                          (4,5) BISPHOSPHATE 5-PHOSPHATASE
 13    (iii) NUMBER OF SEQUENCES: 5
 15     (iv) CORRESPONDENCE ADDRESS:
 16          (A) ADDRESSEE: Incyte Pharmaceuticals, Inc.
 17          (B) STREET: 3174 Porter Drive
 18          (C) CITY: Palo Alto
 19          (D) STATE: CA
 20          (E) COUNTRY: USA
 21          (F) ZIP: 94304
 23      (v) COMPUTER READABLE FORM:
 24          (A) MEDIUM TYPE: Diskette
 25          (B) COMPUTER: IBM Compatible
 26          (C) OPERATING SYSTEM: DOS
 27          (D) SOFTWARE: FastSEQ for Windows Version 2.0
 29     (vi) CURRENT APPLICATION DATA:
C--> 30          (A) APPLICATION NUMBER: US/09/892,287
C--> 31          (B) FILING DATE: 26-Jun-2001
 33    (vii) PRIOR APPLICATION DATA:
 34          (A) APPLICATION NUMBER: 09/258,643
 35          (B) FILING DATE:
 37   (viii) ATTORNEY/AGENT INFORMATION:
 38          (A) NAME: Billings, Lucy J.
 39          (B) REGISTRATION NUMBER: 36,749
 40          (C) REFERENCE/DOCKET NUMBER: PF-0334 US
 42     (ix) TELECOMMUNICATION INFORMATION:
 43          (A) TELEPHONE: 415-855-0555
 44          (B) TELEFAX: 415-845-4166
 46  (2) INFORMATION FOR SEQ ID NO: 1:
 48      (i) SEQUENCE CHARACTERISTICS:
 49          (A) LENGTH: 372 amino acids
 50          (B) TYPE: amino acid
 51          (C) STRANDEDNESS: single
 52          (D) TOPOLOGY: linear
 54    (vii) IMMEDIATE SOURCE:
 55          (A) LIBRARY: BRSTNOT03
 56          (B) CLONE: 638789
```

ENTERED

RAW SEQUENCE LISTING   DATE: 01/24/2002
PATENT APPLICATION: US/09/892,287   TIME: 19:06:31

Input Set : N:\Crf3\RULE60\09892287.raw
Output Set: N:\CRF3\01242002\I892287.raw

```
58     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
60   Met Asp Val Leu Ser Pro Leu Ser Phe Ile Lys Val Ser His Val Arg
61    1               5                  10                  15
62   Met Gln Gly Ile Leu Leu Leu Val Phe Ala Lys Tyr Gln His Leu Pro
63              20                  25                  30
64   Tyr Ile Gln Ile Leu Ser Thr Lys Ser Thr Pro Thr Gly Leu Phe Gly
65          35                  40                  45
66   Tyr Trp Gly Asn Lys Gly Gly Val Asn Ile Cys Leu Lys Leu Tyr Gly
67       50                  55                  60
68   Tyr Tyr Val Ser Ile Ile Asn Cys His Leu Pro Pro His Ile Ser Asn
69   65                  70                  75                  80
70   Asn Tyr Gln Arg Leu Glu His Phe Asp Arg Ile Leu Glu Met Gln Asn
71                  85                  90                  95
72   Cys Glu Gly Arg Asp Ile Pro Asn Ile Leu Asp His Asp Leu Ile Ile
73              100                 105                 110
74   Trp Phe Gly Asp Met Asn Phe Arg Ile Glu Asp Phe Gly Leu His Phe
75              115                 120                 125
76   Val Arg Glu Ser Ile Lys Asn Arg Cys Tyr Gly Gly Leu Trp Glu Lys
77          130                 135                 140
78   Asp Gln Leu Ser Ile Ala Lys Lys His Asp Pro Leu Leu Arg Glu Phe
79   145                 150                 155                 160
80   Gln Glu Gly Arg Leu Leu Phe Pro Pro Thr Tyr Lys Phe Asp Arg Asn
81              165                 170                 175
82   Ser Asn Asp Tyr Asp Thr Ser Glu Lys Lys Arg Lys Pro Ala Trp Thr
83              180                 185                 190
84   Asp Arg Ile Leu Trp Arg Leu Lys Arg Gln Pro Cys Ala Gly Pro Asp
85          195                 200                 205
86   Thr Pro Ile Pro Pro Ala Ser His Phe Ser Leu Ser Leu Arg Gly Tyr
87       210                 215                 220
88   Ser Ser His Met Thr Tyr Gly Ile Ser Asp His Lys Pro Val Ser Gly
89   225                 230                 235                 240
90   Thr Phe Asp Leu Glu Leu Lys Pro Leu Val Ser Ala Pro Leu Ile Val
91              245                 250                 255
92   Leu Met Pro Glu Asp Leu Trp Thr Val Glu Asn Asp Met Met Val Ser
93              260                 265                 270
94   Tyr Ser Ser Thr Ser Asp Phe Pro Ser Ser Pro Trp Asp Trp Ile Gly
95          275                 280                 285
96   Leu Tyr Lys Val Gly Leu Arg Asp Val Asn Asp Tyr Val Ser Tyr Ala
97          290                 295                 300
98   Trp Val Gly Asp Ser Lys Val Ser Cys Ser Asp Asn Leu Asn Gln Val
99   305                 310                 315                 320
100   Tyr Ile Asp Ile Ser Asn Ile Pro Thr Thr Glu Asp Glu Phe Leu Leu
101              325                 330                 335
102   Cys Tyr Tyr Ser Asn Ser Leu Arg Ser Val Val Gly Ile Ser Arg Pro
103              340                 345                 350
104   Phe Gln Ile Pro Pro Gly Ser Leu Arg Glu Asp Pro Leu Gly Glu Ala
105              355                 360                 365
106   Gln Pro Gln Ile
107       370
```

RAW SEQUENCE LISTING   DATE: 01/24/2002
PATENT APPLICATION: US/09/892,287 TIME: 19:06:32

Input Set : N:\Crf3\RULE60\09892287.raw
Output Set: N:\CRF3\01242002\I892287.raw

```
109  (2) INFORMATION FOR SEQ ID NO: 2:
111      (i) SEQUENCE CHARACTERISTICS:
112              (A) LENGTH: 2573 base pairs
113              (B) TYPE: nucleic acid
114              (C) STRANDEDNESS: single
115              (D) TOPOLOGY: linear
117      (vii) IMMEDIATE SOURCE:
118              (A) LIBRARY: BRSTNOT03
119              (B) CLONE: 638789
121      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
123  GAAGGCTCAG CATACACGTC GTGACTTGGA ACGTGGCTTC GGCAGCGCCC CTCGAGCTCT    60
124  CAGTGACCTG CTTCAGCTGA ACAACCGGAA CCTCAATCTT GACATATATG TTATTGGTTT   120
125  GCAGGAATTG AACTCTGGGA TCATAAGCCT CCTTTCCGAT GCTGCCTTTA ATGACTCGTG   180
126  GAGCAGTTTC CTCATGGATG TGCTTTCCCC TCTGAGCTTC ATCAAGGTCT CCCATGTCCG   240
127  TATGCAGGGG ATCCTCTTAC TGGTCTTTGC CAAGTATCAG CATTTGCCCT ATATCCAGAT   300
128  TCTGTCTACT AAATCCACCC CCACTGGCCT GTTTGGGTAC TGGGGGAACA AAGGTGGAGT   360
129  CAACATCTGC CTGAAGCTTT ATGGCTACTA TGTCAGCATC ATCAACTGCC ACCTGCCTCC   420
130  CCACATTTCC AACAATTACC AGCGGCTGGA GCACTTTGAC CGGATCCTGG AGATGCAGAA   480
131  TTGTGAGGGG CGAGACATCC CAAACATCCT GGACCACGAC CTCATTATCT GGTTTGGAGA   540
132  CATGAACTTT CGGATCGAGG ACTTTGGGTT GCACTTTGTT CGGGAATCCA TTAAAAATCG   600
133  GTGCTACGGT GGCCTGTGGG AGAAGGACCA GCTCAGCATT GCCAAGAAAC ATGACCCGCT   660
134  GCTCCGGGAG TTCCAGGAGG GCCGCCTACT CTTCCCGCCC ACCTACAAGT TTGATAGGAA   720
135  CTCCAACGAC TATGACACCA GTGAGAAAAA ACGCAAGCCT GCATGGACCG ATCGCATCCT   780
136  GTGGAGGCTG AAGCGGCAGC CCTGTGCTGG CCCCGACACT CCCATACCGC CGGCGTCACA   840
137  CTTCTCCTTG TCTCTGAGGG GCTACAGCAG CCACATGACG TACGGCATCA GCGACCACAA   900
138  GCCTGTCTCC GGCACGTTCG ACTTGGAGCT GAAGCCATTG GTGTCTGCTC CGCTGATCGT   960
139  CCTGATGCCC GAGGACCTGT GGACCGTGGA AAATGACATG ATGGTCAGCT ACTCTTCAAC  1020
140  CTCGGACTTC CCCAGCAGCC CGTGGGACTG GATTGGACTG TACAAGGTGG GGCTGCGGGA  1080
141  CGTTAATGAC TACGTGTCCT ATGCCTGGGT CGGGGACAGC AAGGTCTCCT GCAGCGACAA  1140
142  CCTGAACCAG GTTTACATCG ACATCAGCAA TATCCCTACC ACTGAAGATG AGTTTCTCCT  1200
143  CTGTTACTAC AGCAACAGTC TGCGTTCTGT GGTGGGGATA AGCAGACCCT TCCAGATCCC  1260
144  GCCTGGCTCC TTGAGGGAGG ACCCACTGGG TGAAGCACAG CCACAGATCT GAGCCAGGAT  1320
145  GGGAGTGAAT CCCAGGCGGA GGCCAGAGCT GGCAGCCAGC TCTGCCTTTC CACTGCCGGG  1380
146  AGTGCTGGGG GCCCAGCCTG GCCCCCTGAA GAGACAGCCA AGTGTCGTCC ACATACTCCT  1440
147  CCCAGAGTGA GCTCTAACCA GGCTCATTTG CTCTCTCCAC TACTCATCTC TGGAATTAGC  1500
148  CGCTTAAATA CAGGTTTTTG TTGCTGAGAT GTGAGTGAAA CCAGCTAGTG TGTCAACAGT  1560
149  GAAGACCTGG GGACAGTTCT GCGTCTCATT TCTGGATTCC TACCCCCTCT TCTAGTCTTG  1620
150  CCCAAGTAGT CCTGCCAGGC ACATGCCCCA TTTGGCACAG GCCTGCATTC TTGTCGTGCC  1680
151  GTCCTGGGCC TCAGGCTGTC TGGGAGGGGA GATGCTCACA TTTGTACAGG CTACATAGAC  1740
152  TGGTGCAAGC AGTGCTGGAT TCCAGGAGTC TTGGCATCTC ATAGCTTGTC CCCGTGAGGA  1800
153  GTGAGCAGAG GGTCTGGGAT TTCTGCTTTC AGCAAAAGCA GTCTGACTCA GTGGGCAGAA  1860
154  TGGAGGGGCC CCTCTAGCCA GGCTCTTACG CCATGGTTAT GAGCAGGTTG ATGAGGGTCC  1920
155  TTCGGCCAGC ACAACCTTCC TCCCTACTCA CGGCATGGAG TCTGACTGCA TGGAAGTTCC  1980
156  AGATCCTGAC AGAGAGAACT GGGAAGGATC CAGGTTCGCT TCCGTTGGTA GCTTGAGTCC  2040
157  CATGCCTCCA CCCTGCCATC TGAGGAAGGG GTGACAAGTG GTCAAGGAGC TGTGGCCACA  2100
158  GACTTTTCCA GGGTGGTCCT TGGCAGGTGA GGTGCGTCTG TGCCACCCTT GTCAGGAGTC  2160
159  ATTGACGACG GCCCCCCCCT GGACCCCCCG GGACCTCAGA GTGGGGGCAG GCAGAAGGGA  2220
160  GAACCAGCTC AAGACATTTT GGAGGATCTG GCCCTGGGGT TCTTCAGAGA ACACCCTCTA  2280
161  GGGGCTTTGG GGACATGGCC TGTCCCCACA TCCAGCACTT GCCTCCGCCA TGGTCACTCG  2340
```

RAW SEQUENCE LISTING   DATE: 01/24/2002
PATENT APPLICATION: US/09/892,287   TIME: 19:06:32

Input Set : N:\Crf3\RULE60\09892287.raw
Output Set.: N:\CRF3\01242002\I892287.raw

```
162  GCAGCCCTTT TCCCAGGAGA AGACACCTCT GGGAGCCTGC TCAGTGCTTG TCCTGCCATC  2400
163  CTGTGTCCTG GGACTGAGGG TTACTCCAGT TGCTCTGTGT TGCATACTCT CCCCCGCAAG  2460
164  CCTGTGTATG AAGAATTGTC CCCTGGCTTC CAGCAGGCCA TGGCTGGCTG TTTTGTGACT  2520
165  GTTACATTGT GCAGGGGTAA TTATTAGCGT GGCTTTTACA CTTAAAAAAA AAA         2573
167  (2) INFORMATION FOR SEQ ID NO: 3:
169      (i) SEQUENCE CHARACTERISTICS:
170          (A)  LENGTH: 329 amino acids
171          (B)  TYPE: amino acid
172          (C)  STRANDEDNESS: single
173          (D)  TOPOLOGY: linear.
175     (vii) IMMEDIATE SOURCE:
176          (A)  LIBRARY: GenBank
177          (B)  CLONE: 1399101
179      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
181  Ala Gly Leu Phe Gly Tyr Trp Gly Asn Lys Gly Gly Val Asn Ile Cys
182  1                   5                   10                  15
184  Leu Lys Leu Tyr Gly Tyr Tyr Val Ser Ile Ile Asn Cys His Leu Pro
185              20                  25                  30
186  Pro His Ile Ser Asn Asn Tyr Gln Arg Leu Glu His Phe Asp Arg Ile
187          35                  40                  45
188  Leu Glu Met Gln Asn Cys Glu Gly Arg Asp Ile Pro Asn Ile Leu Asp
189      50                  55                  60
190  His Asp Leu Ile Ile Trp Phe Gly Asp Met Asn Phe Arg Ile Glu Asp
191  65                  70                  75                  80
192  Phe Gly Leu His Phe Val Arg Glu Ser Ile Lys Asn Arg Cys Tyr Gly
193              85                  90                  95
194  Gly Leu Trp Glu Lys Asp Gln Leu Ser Ile Ala Lys Lys His Asp Pro
195              100                 105                 110
196  Leu Leu Arg Glu Phe Gln Glu Gly Arg Leu Leu Phe Pro Pro Thr Tyr
197          115                 120                 125
198  Lys Phe Asp Arg Asn Ser Asn Asp Tyr Asp Thr Ser Glu Lys Lys Arg
199      130                 135                 140
200  Lys Pro Ala Trp Thr Asp Arg Ile Leu Trp Arg Leu Lys Arg Gln Pro
201  145                 150                 155                 160
202  Cys Ala Gly Pro Asp Thr Pro Ile Pro Pro Ala Ser His Phe Ser Leu
203              165                 170                 175
204  Ser Leu Arg Gly Tyr Ser Ser His Met Thr Tyr Gly Ile Ser Asp His
205              180                 185                 190
206  Lys Pro Val Ser Gly Thr Phe Asp Leu Glu Leu Lys Pro Leu Val Ser
207          195                 200                 205
208  Ala Pro Leu Ile Val Leu Met Pro Glu Asp Leu Trp Thr Val Glu Asn
209      210                 215                 220
210  Asp Met Met Val Ser Tyr Ser Ser Thr Ser Asp Phe Pro Ser Ser Pro
211  225                 230                 235                 240
212  Trp Asp Trp Ile Gly Leu Tyr Lys Val Gly Leu Arg Asp Val Asn Asp
213              245                 250                 255
214  Tyr Val Ser Tyr Ala Trp Val Gly Asp Ser Lys Val Ser Cys Ser Asp
215              260                 265                 270
216  Asn Leu Asn Gln Val Tyr Ile Asp Ile Ser Asn Ile Pro Thr Thr Glu
```

RAW SEQUENCE LISTING   DATE: 01/24/2002
PATENT APPLICATION: US/09/892,287   TIME: 19:06:32

Input Set : N:\Crf3\RULE60\09892287.raw
Output Set: N:\CRF3\01242002\I892287.raw

```
217              275.          .   280                 285
218   Asp Glu Phe Leu Leu Cys Tyr Tyr Arg Asn Ser Leu Arg Ser Val Val
219        290                   295               300
220   Gly Ile Arg Arg Pro Phe Gln Ile Pro Pro Gly Ser Leu Arg Glu Asp
221   305                   310               315               320
222   Pro Leu Gly Glu Ala Gln Pro Gln Ile
223               · 325
225   (2) INFORMATION FOR SEQ ID NO: 4:
227       (i) SEQUENCE CHARACTERISTICS:
228           (A) LENGTH: 942 amino acids
229           (B) TYPE: amino acid
230           (C) STRANDEDNESS: single
231           (D) TOPOLOGY: linear
233      (vii) IMMEDIATE SOURCE:
234           (A) LIBRARY: GenBank
235           (B) CLONE: 1019103
237.      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
239   Val Thr Val Pro Glu Pro Gly Ala Ala Glu Ser Arg Ala Pro Cys Gly
240    1               5                   10                15
241   Asp Ser Ser Gly Gly Cys Val Arg Ser Ala Gly Ala Ser Met Asp Gln
242           20                   25                30
243   Ser Val Ala Ile Gln Glu Thr Leu Ala Glu Gly Glu Tyr Cys Val Ile
244           35                   40                45
245   Ala Val Gln Gly Val Leu Cys Glu Gly Asp Ser Arg Gln Ser Arg Leu
246    50                   55                60
247   Leu Gly Leu Val Arg Tyr Arg Leu Glu His Gly Gly Gln Glu His Ala
248   65                   70                75                80
249   Leu Phe Leu Tyr Thr His Arg Arg Met Ala Ile Thr Gly Asp Asp Val
250                ·85                  · 90               95
251   Ser Leu Asp Gln Ile Val Pro Val Ser Arg Asp Phe Thr Leu Glu Glu
252           100                  105               110
253   Val Ser Pro Asp Gly Glu Leu Tyr Ile Leu Gly Ser Asp Val Thr Val
254           115                  120               125
255   Gln Leu Asp Thr Ala Glu Leu Ser Leu Val Phe Gln Leu Pro Phe Gly
256        130                  135               140
257   Ser Gln Thr Arg Met Phe Leu His Glu Val Ala Arg Ala Cys Pro Gly
258   145                  150              \ 155               160
259   Phe Asp Ser Ala Thr Arg Asp Pro Glu Phe Leu Trp Leu Ser Arg Tyr
260               165                  170               175
261   Arg Cys Ala Glu Leu Glu Leu Glu Met Pro Thr Pro Arg Gly Cys Asn
262               180                  185               190
263   Ser Ala Leu Val Thr Trp Pro Gly Tyr Ala Thr Ile Gly Gly Gly Gly
264           195                  200               205
265   Ser Asn Phe Asp Gly Leu Arg Pro Asn Gly Lys Gly Val Pro Met Asp
266        210                  215               220
267   Gln Ser Ser Arg Gly Gln Asp Lys Pro Glu Ser Leu Gln Pro Arg Gln
268   225                  230               235               240
269   Asn Lys Ser Lys Ser Glu Ile Thr Asp Met Val Arg Ser Ser Thr Ile
270               245                  250               255
```

VERIFICATION SUMMARY   DATE: 01/24/2002
PATENT APPLICATION: **US/09/892,287**   TIME: 19:06:33

Input Set : **N:\Crf3\RULE60\09892287.raw**
Output Set: **N:\CRF3\01242002\I892287.raw**

L:30 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:31 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]